Cluster Virtualization

Team Members: Russell Husted, Michigan Technological University Derek Walker, North Carolina Agriculture & Technical State University
Povi L. Cruz, Northern New Mexico College

Mentors: Jharrod LaFon (HPC-3), James Williams (HPC-3)

Los Alamos NATIONAL LABORATORY EST.1943

Abstract

Every year, Los Alamos National Laboratory (LANL) scientists create numerous small clusters for their research computing needs. This research does not require a full supercomputer but calls for more power than a workstation can provide. This project explores the cloud computing abstraction enabled by Eucalyptus. Eucalyptus is a set of software that allows dynamic allocation and isolation of virtual machine groups. Scientists can then create and use relatively small clusters without purchasing additional hardware. The goal of this project is to provide an internal system functionally identical to Amazon's EC2 cloud service, which allows the creation, management and modification of virtual systems. National security concerns prevent the use of commercially available cloud computing services.

Motivation

Future uses of this project include an Amazon cloud-like system which would be implemented inside LANL's secured network and providing a centralized source for scientists to create small computing clusters. The Amazon cloud cannot be used for security purposes since it allows foreign nationals to have root on these machines, allowing for greater research and development. The set up utilizes the Amazon API, existing software should be compatible. Another benefit of this project is the simplicity & convenience by allowing the scientist to perform research without the burden of cluster management.

PRObE Parallel Reconfigurable Observational Environment

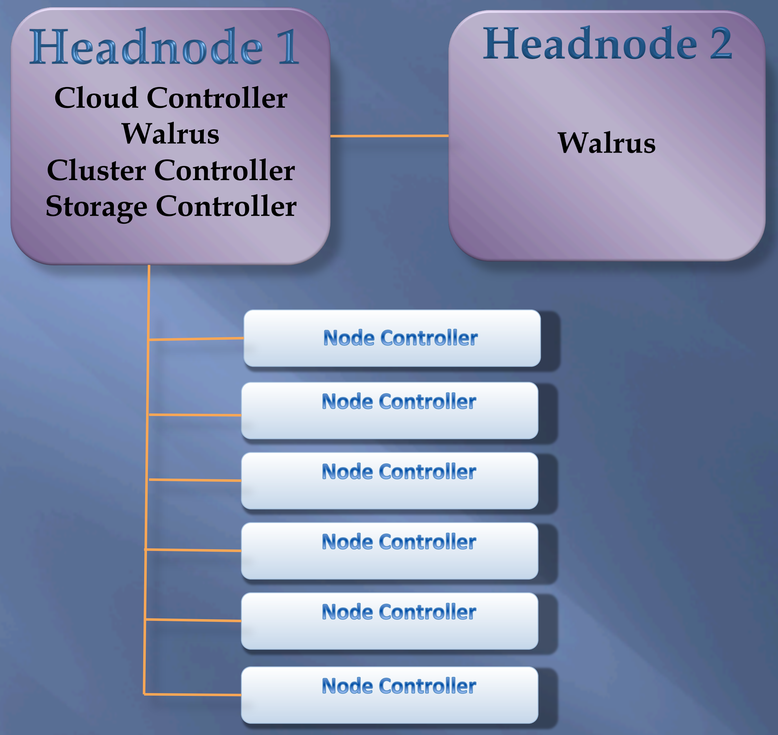
National Security Education Center A CONSORTIUM OF LANL INSTITUTES

NSF

ISTI INFORMATION SCIENCE & TECHNOLOGY INSTITUTE

New Mexico CONSORTIUM NM STATE UNIVERSITY NEW MEXICO TECH The University of New Mexico

Architecture



Eucalyptus – Elastic Utility Computing Architecture Linking Your Programs to Useful Systems

Cloud Controller: Main entry-point to the cloud, determines high level scheduling then implements using cluster controllers

Cluster Controller: Manages the virtual instance network by gathering information about a set of VMs and schedules VM execution on designated NCs

Node Controller: Controls execution, inspection and termination of VM instances. Also manages local copies of instance images and queries and controls the system software.

Storage Controller: Implements block-accessed network storage and adapts to interfacing with various storage systems

Walrus (put/get storage): Enables users to store data, organized later as consistent buckets and objects

Cluster Specification

Hardware:
Head nodes (2) – Dual Intel ® Xeon ® CPU E5410, 2.33GHz. Quad core CPUs each with 6MB L2 Cache. 8GB of DDR2-800MHz RAM, an 80GB SATA hard drive, two gigabit Ethernet ports, two 10 Gigabit Ethernet ports.

Xen™

Compute Nodes (6) – Dual Nahalem Intel ® Xeon ® CPU E5504, running at 2.00GHz. Quad core CPUs equipped with 512 kB of private L2 cache for each core and shared 4MB L3 cache

CentOS

Software:
Eucalyptus (v2.0)
Xen (v3.0)
CentOS (v5.6)

Eucalyptus Systems

Moving Forward

- Analysis of the affect of virtualization on performance
- IO performance is a major concern when virtualizing
- Testing and analysis at larger scale (larger number of nodes).
- Testing on different types of underlying hardware
- Implementation of a full production system

LA-UR-11-111087